Problem Statement: Bank of Portugal Telephone Campaign Predicting Model

Jess Zhang

April 9, 2018

Project Background

Bank of Portugal uses telephone campaign to sell term deposits is a annual event organized by the internal campaign team. Each telephone call is costly. In order to improve its telemarketing effeciency, the campaign team in Bank of Portugal wants to be able to predict the campaign outcome.

The team has collected past campaign data and stored the information in a csv file.

Stakeholders

The key players involved in this project are:

- John Smith – Campaign Manager from Bank of Portugal's Telemarketing department. He is the main stakeholder and point of contact for Bank of Portugal. John and his team will be using the model to analyze their client base and develop campaign strategy accordingly.
  Email: John.Smith@bancdeportugal.com

- Ashley Harvey – Data Analyst from Bank of Portual. She works on John Smith's team. Her job is to collect and record client and campaign data. Once the model is built, Ashley will be the primary user of this model to report results to John. Email: Ashley.Harvey@bancdeportugal.com

- Jessica Zhang – Data Scientist from CSX415. She is the assigned to this project to develop an effective machine learning model that can help Bank of Portugal to understand and improve their campaign effectiveness. Email: Jesszh@Gmail.com

- Nik Nolasco – Data Engineer from CS415. He is assigned to this project to analyze and clean the dataset provided by Bank of Portugal and prepare the dataset to a clean state so the data scientist can analyze and train the data effeciently. Email: Nick.N@gmail.com

Project Scope

Many factors, such as age, job, and loan amount can affect people's banking decisions. Bank of Portugal believes it needs to develop a better targeted marketing strategy towards different profile groups. The goal of this project is to provide the campaign department with meaningful grouping of its clients from past campaigning data.

The current campaign success rate from this dataset is at 23.6%. In order to improve the success rate, the campaign manager needs to be able to first "predict" the success likelihood bofore he can direct his team with appropriate actions. Since there is no pre-existing model to compare with, the CSX415 data science team will use a "naive analysis" as the baseline.

The model we have build should be able to predict at 80 percent accuracy to be useful.

Developing of the marketing strategy is outside of the scope of this project.

A few algorithms will be run and the best suited model will be built to analyze the given dataset. The analysis will be shared and discussed with the campaign manager of Bank of Portugal. The campaign manager and his team can continue to use the model to analyze data in the future.

The delivered model will work with additional data in the dataset. However, the model may not work if the client adds or removes columns from the dataset without conselingn this project team.

Benefits of the Project

With a model which can predict at least 80 percent of the outcome, the campaign manager can focus most his team's effort on the clients who are most likely to make a term deposit during busy season and call the rest during low season.

The campaign team can also develop different call tactic towards different profile group therefore improve the overall campaign effectiveness.

Witht future data gathering, the prediction accuracy of the model will continue to improve. The campaign manager can futhrer improve his strategy.

Estimated Risks

It is reasonable to believe that there are such groups exist in the given dataset. It is unlikely however possible that no meaningful profile groups can be identified from given data.

Delivery Timeline

The project team will be working on this project for the next 2 months. One data scientist and one data engineer will be working on the model and analysis full-time.

Milestone 1: 4/30/2018 The preliminary result will be delivered and discussed with the client on 4/30/2018.

Final Delivery: 5/31/2018 *The estimated final delivery date is May 31st, 2018. A final demo will be given by CSX415 to the campaign team from Bank of Portugal.* The final delivery date will be dicsussed with the client and adjusted as needed at the 4/30/2018 milestone date. *The CSX415 team will train one representative data analyst from Bank of portugal on the usage and the basic maintenance of the model.

Client Responsibilities After Delivery

Once deployed, the data model should be run and reviewed as needed before the campaign planning starts.

New campaign data should be updated and the model should be retrained as the dataset changes.

Do not change the format of the data or the model may not run again!

Since the users of this model may not have sufficient knowledge about machine learning, it is possible for someone to accidentally remove or damage the model or one of the supporting files. Weekly backup of the data as well as all files from this deliverable is highly recommended.

The data model's performance will improve with more data gathering and trainging. However, after a period of time, the new data may be different from the orginial data enough that the difference will affect the model performance negatively. CSX41 recommends Bank of Portugal to bring the model back for performance tuning once the model performance decrases more than 10 percent, or annually, whichever comes first.

Data Source

Bank Marketing Data from http://archive.ics.uci.edu/ml/datasets/Bank+Marketing